DESCRIPTION

METHOD FOR PRODUCING YELLOW FLOWER BY CONTROLLING FLAVONOID SYNTHETIC PATHWAY

Technical Field

The present invention relates to a gene coding for a protein having activity of transferring sugars to chalcones, and to plants with modified flower color utilizing the gene. More specifically, the invention relates to a gene coding for a protein having activity of synthesizing chalcone 4'-glucoside, and preferably to a gene derived from the family *Scrophulariaceae*, and more preferably *Antirrhinum majus* or *Linaria bipartita*, coding for a protein having activity of synthesizing chalcone 4'-glucosides, as well as to a method for modifying flower color, and preferably a method for modifying flower color to yellow, by expressing these genes and an aureusidin synthase (hereinafter, "AS") gene either separately or simultaneously and accumulating chalcones or aurones.

Background Art

Flower color is an important feature for the appreciation and purchasing of ornamental flowers, and flowers with a large variety of colors have traditionally been bred. It is rare for a single species to possess flowers of all colors, as the biosynthesis of pigments that appear as flower colors is genetically determined. Because the gene sources that can be used in hybridization breeding are limited to crossable related varieties, it is essentially impossible to produce flowers of all colors in a target variety by hybridization breeding. Recently, gene recombination techniques have made it possible to obtain flower pigment-synthesizing genes from certain plants and express those genes in different species in order to

achieve modified flower color (Plant Cell Physiol. 39, 1119(1998), Curr. Opin. Biotechnol. 12, 155 (2001)).

The flower colors of orange, red, violet and blue are exhibited primarily by flavonoids known as anthocyanins. Yellow colors generally derive from non-flavonoid compounds such as carotenoids and betalains, but the yellow colors of some plant species are due to flavonoids. For example, yellow carnations are known to possess 4,2',4',6'-tetrahydroxychalcone (hereinafter, THC) 2'-glucoside in their flower petals (Phytochemistry 5, 111 (1996)). THC 4'-glucoside is also found in *Antirrhinum majus* and *Linaria bipartita*.

Chalcones such as, THC, butein, isoliquiritigenin and their glycosilated derivatives are known; for example, the aglycon of the glucosides in carnations, morning glory, peony, aster, strawflower, periwinkle, cyclamen and petunia is THC, in *Antirrhinum majus* (snapdragon) and statice it is 3,4,2',4',6'-pentahydroxychalcone (PHC), in cosmos and Jerusalem artichoke it is butein, and in dahlia it is butein and isoliquiritigenin. Also, certain limited species such as snapdragon, *Linaria bipartita* (toadflax) and morning glory contain yellow flower pigments known as aurones, including aureusidin (hereinafter, AU) and bracteatin.

Because the absorption maxima for aurones are between 399 and 403 nm, compared to absorption maxima between 372 and 382 nm for chalcones, their color tones differ and the fluorescence emitted gives aurones a sharper yellow color (Biohorti 1, 49-57 (1990), Seibundo Shinkosha). Chalcones, aurones and anthocyanins usually accumulate as glucosides in the vacuoles of plant cells. The biosynthetic pathway of anthocyanins has been thoroughly studied, and the enzymes involved in anthocyanin synthesis and their coding genes are known (Comprehensive Natural Products Chemistry, vol I (ed. Sankawa) pp713-748, Elsevier, Amsterdam (1999)).

The biosynthetic pathway of flavonoids is widely

distributed among the higher plants and is conserved among species. THC is biosynthesized from three molecules of malonyl CoA and one molecule of coumaroyl CoA, by the catalytic action of chalcone synthase. THC exhibits a light yellow color, but in plant cells it is usually rapidly converted to colorless naringenin by chalcone isomerase (CHI). Also, THC is highly unstable at near neutral pH and is converted to naringenin by spontaneous ring closure. For THC to exist stably in plant cells, i.e. for it to stably exhibit a yellow color, the 2'-position of THC must be modified with a saccharide to prevent its ring closure. The reaction is catalyzed by an enzyme that transfers glucose to the 2'-position of THC (UDP-glucose: 4,2',4',6'-tetrahydroxychalcone 2'-glucosyltransferase, hereinafter abbreviated as 2'CGT).

THC 2'-glucoside is present in carnation and cyclamen, and therefore 2'CGT is also predicted to be found in their flowers. Thus, it was conjectured that if the 2'CGT gene could be obtained and the enzyme gene expressed in a plant, it should be possible to accumulate THC 2'-glucosides and produce yellow flowers (Biotechnology of Ornamental Plants, Edited by Geneve, Preece and Merkle, pp259-294, CAB International Wallingford, UK (1997)). Moreover, it was discovered that adequately accumulating THC 2'-glucoside and exhibiting yellow color requires deletion of the CHI gene to suppress enzymatic conversion from THC to naringenin, and that a clear yellow color also requires deletion of the gene for flavanone 3-hydroxylase (hereinafter abbreviated as F3H) in addition to the CHI gene (Plant Cell Physiol. 43, 578 (2002)).

While cloning of the carnation 2'CGT gene has been reported to date (Plant Cell Physiol. 44, s158 (2003)), its sequence has not been published. Also, the gene coding for 2'CGT activity has been obtained from carnation and expressed in petunia, thereby accumulating

THC 2'-glucoside in petunia petals (PCT/JP03/10500). However, the THC 2'-glucoside produced by 2'CGT does not have a chemical structure that can serve as a precursor for aurone synthesis. Also, as mentioned above, accumulation of THC 2'-glucoside results in only light yellow petals.

It is known that faint yellow petals are produced by accumulation of THC having the 2'-hydroxyl methylated, but the nature of the enzyme that catalyzes this methylation and of its gene is unknown. Yellow varieties such as dahlia and cosmos contain 6'-deoxychalcone. In legumes, 6'-deoxychalcone is the precursor of 5-deoxyflavonoid, which is synthesized by the catalytic action of chalcone synthase (CHS) and chalcone reductase (CHR). It has been reported that introduction of the alfalfa CHR gene into petunia produced 6'-deoxychalcones such as butein, and that when the CHR gene was introduced into white flower petunia the flowers were mostly white upon blooming although a very light yellow color was observed at the budding stage, and therefore it was not possible to create an industrially useful yellow flower (Plant J. 13, 259 (1998)).

Because aurones exhibit a more brilliant yellow color than chalcone glucosides as explained above, it would be highly useful, for the industry, to develop a method for accumulating aurones. AS, one of the enzymes involved in aurone biosynthesis, and its gene, have already been reported in the literature (Science, 290, 1163 (2000)). According to this report, AS produces AU, bracteatin and their glucosides from THC, PHC and their glucosides as substrates. However, use of the AS gene to produce and accumulate aurones such as AU and bracteatin has not been described.

The present inventors have constructed a binary vector having the AS gene linked downstream from a structural promoter and introduced the AS gene into petunia and torenia by the *Agrobacterium* method, but no

accumulation of aurones was observed. It has also been reported that 3-glucosylation of anthocyanidin is essential for transportation of anthocyanins into vacuoles (Nature 375, 397 (1995)), suggesting that glucosylation of aurones is likewise necessary as a transportation signal into vacuoles. In fact, the major aurone that accumulates in yellow *Antirrhinum majus* flower petals is AU 6'-glucoside. A GT exhibiting AU 6'-glucosylating activity (AU6GT) (WO 00/49155) was therefore obtained, and the AU6GT gene was constitutively expressed in petunia together with the AS gene, but accumulation of aurones was not observed.

The enzymes involved in flavonoid and anthocyanin biosynthesis are believed to localize in the cytoplasm or endoplasmic reticulum of cells. By the actions of these enzymes, flavonoids and anthocyanins are synthesized and glucosylated outside of vacuoles, i.e. in the cytoplasm, and transported into the vacuoles (Natural Product Reports 20, 288, (2003)). However, ardent research led the present inventors to the finding that AS is exceptional in that it localizes in the vacuoles. Thus, it was hypothesized that glucosylated chalcones may be transported to the vacuoles *in vivo* and used as substrates for synthesis of aurones in the vacuoles.

As mentioned above, AU 6'-glucoside is the major aurone that accumulates in the vacuoles of yellow *Antirrhinum majus* petals. The 6' position of AU corresponds to the 4' position of THC, and THC 4'-glucosides are also present in yellow *Antirrhinum majus* petals. On this basis, an aurone synthetic pathway was inferred wherein glucosylation of the 4' position of THC synthesized in the cytoplasm is followed by transport to the vacuoles, and this is used as substrate for synthesis of AU 6'-glucosides by AS. Thus, it was concluded that synthesis of THC 4'-glucoside is essential for synthesis of aurones such as AU 6'-glucoside in different plant varieties. For this purpose, UDP-glucose:4,2',4',6'-

tetrahydroxychalcone 4'-glucosyltransferase (hereinafter, 4'CGT) is required for 4'-glucosylation of THC, and therefore the 4'CGT gene must be obtained. However, cloning of the 4'CGT gene has not yet been reported, nor do reports exist of isolating 4'CGT.

Enzymes that catalyze glucosylation of a variety of compounds including flavonoids to produce glucosides are generally referred to as glucosyltransferases (GT), and plants possess a large diversity of GT molecules and their coding genes, corresponding to the types of substrates and transferred sugars. Because GT enzymes usually utilize UDP-glucose as the glucose donor, they contain in their amino acid sequence a motif that binds UDP-glucose (Plant Physiol. 112, 446 (2001)). Already, GT genes carrying this motif are known in 99 species of *Arabidopsis* whose entire genome structure has been elucidated (J. Biol. Chem. 276, 4338, (2001)).

GT enzymes and amino acid sequences and functions have also been worked out in several other plants. The genes for enzymes catalyzing reactions of transferring sugars to the 3-hydroxyl groups of flavonoids or anthocyanidins (UDP-glucose:flavonoid 3-glucosyltransferase, hereinafter: 3GT) have been obtained from perilla, corn, gentian, grape and the like (J. Biol. Chem. 274 , 7405 (1999); J. Biol. Chem. 276, 4338, (2001)). In addition, genes for enzymes catalyzing reactions of transferring sugars to the 5-hydroxyl groups of anthocyanins (UDP-glucose:anthocyanin 5-glucosyltransferase, hereinafter: 5GT) have been obtained from perilla, verbena and the like (J. Biol. Chem., 274, 7405, (1999)).

Analysis of the amino acid sequences of 3GT and 5GT has shown that GT enzymes with the same function have similar amino acid sequences even in different plant varieties, or in other words, that they constitute a family (J. Biol. Chem. 276, 4338, (2001)). Thus, it is not difficult to obtain enzymes having the same function

as known GT enzymes (i.e., orthologs) from other plant varieties, given the current level of technology. For example, the petunia 5GT gene has been cloned using the perilla 5GT gene (Plant Mol Biol. 48, 401 (2002)). However, much laborious trial and error is required to obtain a novel GT gene having absolutely no known ortholog.

As regards *Arabidopsis* whose entire genome structure is known as mentioned above, its flower petals are white and accumulation of chalcone 4'-glucosides has not been reported. Consequently, *Arabidopsis* GT gene information cannot be used for cloning of the 4'CGT gene. Moreover, even though 2'CGT has been isolated from carnation (PCT/JP03/10500), high homology does not necessarily exist between the 4'CGT gene and the 2'CGT gene. This is because the biochemical and molecular biological features of each GT may differ substantially if the position of sugar addition is different, even if the substrate is the same. This is also supported by the fact that 3GT and 5GT belong to different GT families. Also, betanidine 5GT and 6GT have the same substrates, and yet their amino acid homology has been reported to be only 19% (Planta 214, 492 (2002)).

In fact, the GT enzymes that transfer sugars to the 3-, 5- and 3'-positions of the same anthocyanidin skeleton belong to different families in the GT superfamily, and the amino acid homology between these families is no more than about 20% (Plant Physiol. 132, 1652, (2003), Natural Product Reports 20, 288, (2003)). Several methods are possible for obtaining not only the 4'CGT gene but also novel genes. For example, the genes for enzymes expressed in flower petals, which are involved in the synthesis of rose scent components, have been extensively sequenced and have been identified by their structures, expression patterns and expression in *E. coli* (Plant Cell. 14, 2325 (2002)). In order to identify the 4'CGT gene, 5,000 clones were randomly

selected from a cDNA library derived from the petals of yellow *Antirrhinum majus* (Butterfly Yellow variety) which accumulates aurone and chalcone 4'-glucosides, and their nucleotide sequences were determined.

As a result of homology search using a public DNA database, three different GT genes were obtained. Two of the genes were the 3GT gene and the aforementioned AU6GT-coding gene (WO 00/49155), while the remaining gene was for a novel GT (designated as pSPB662) (SEQ ID NO: 13). However, the GT encoded by pSPB662 exhibited no glucosylating activity for THC, and was clearly not 4'CGT. Furthermore, as mentioned above, high expression of this gene together with the AS gene in petunia resulted in no observable production of chalcone glucosides or aurones, nor was any change in flower color seen. These results suggest that a chalcone glucosylating enzyme gene cannot be isolated by random screening of approximately 5,000 clones, and it was therefore difficult to obtain a 4'CGT gene.

Patent document 1: PCT/JP03/10500

Patent document 2: WO 00/49155

Non-patent document 1: Plant Cell Physiol. 39,1119 (1998)

Non-patent document 2: Curr. Opin. Biotechnol. 12, 155 (2001)

Non-patent document 3: Phytochemistry 5,111 (1996)

Non-patent document 4: Biohorti 1 49-57 (1990) SEIBUNDO SHINKOSHA

Non-patent document 5: Comprehensive Natural Products Chemistry, vol I (ed. Sankawa) pp713-748, Elsevier, Amsterdam (1999)

Non-patent document 6: Biotechnology of Ornamental Plants, Edited by Geneve, Preece and Merkle, pp259-294, CAB International Wallingford, UK (1997)

Non-patent document 7: Plant Cell Physiol. 43, 578 (2002)

Non-patent document 8: Plant Cell Physiol. 44, s158

(2003)
Non-patent document 9: Plant J. 13, 259 (1998)
Non-patent document 10: Science, 290, 1163 (2000)
Non-patent document 11: Nature 375, 397 (1995)
Non-patent document 12: Natural Product Reports 20, 288, (2003)
Non-patent document 13: Plant Physiol. 112, 446 (2001)
Non-patent document 14: J. Biol. Chem. 276, 4338, (2001)
Non-patent document 15: J. Biol. Chem. 274, 7405 (1999)
Non-patent document 16: Plant Mol Biol. 48, 401 (2002)
Non-patent document 17: Planta 214, 492 (2002)
Non-patent document 18: Plant Physiol. 132, 1652, 2003 (2003)
Non-patent document 19: Plant Cell. 14, 2325 (2002)

Disclosure of the Invention

The present invention provides a protein having activity of transferring sugars to the 4'-hydroxyl groups of chalcones, and its gene, and preferably a protein having activity of transferring a sugar to the 4'-hydroxyl groups of chalcones specifically, and its gene. The invention further provides plants with modified flower color, preferably altered to yellow, using the aforementioned GT gene.

As mentioned above, the biochemical and molecular biological features of 4'CGT are unknown, and the enzyme has not been purified nor its gene cloned. The present inventors, using a probe having a nucleotide sequence corresponding to the conserved amino acid sequence of the GT family, obtained ten different GT genes having the nucleotide sequence for the conserved amino acid sequence from a yellow *Antirrhinum majus* (Butterfly Yellow) flower petal cDNA library. The GT gene group was expressed in

different *E. coli* cells, activity of transferring glucose to the chalcone 4'-position, i.e. 4'CGT activity, was confirmed among the *E. coli* extracts, and it was confirmed that the cloned gene coded for 4'CGT. The gene was expressed in plants for modification of flower color, and the present invention was thereupon completed.

Specifically, the invention provides (1) a gene coding for a protein having activity of transferring a sugar to the chalcone 4'-position.

The invention further provides (2) a gene according to (1) above coding for the amino acid sequence listed as SEQ ID NO: 2.

The invention further provides (3) a gene according to (1) above, which hybridizes to all or a portion of the nucleotide sequence listed as SEQ ID NO: 1 under conditions of 5 x SSC, 50°C and codes for a protein having activity of transferring a sugar to the chalcone 4'-position.

The invention still further provides (4) a gene according to (1) above, which codes for a protein having the amino acid sequence listed as SEQ ID NO: 2 with a modification of one or a plurality of amino acids that are added, deleted and/or substituted with other amino acids, and having activity of transferring a sugar to the chalcone 4'-position.

The invention still further provides (5) a gene according to (1) above, which hybridizes to DNA comprising all or a portion of the nucleotide sequence listed as SEQ ID NO: 1 under stringent conditions and codes for a protein having activity of transferring a sugar to the chalcone 4'-position.

The invention still further provides (6) a gene according to any one of (1) to (5) above, which is derived from the family *Scrophulariaceae*.

The invention still further provides (7) a vector comprising a gene according to any one of (1) to (6) above.

The invention still further provides (8) host cells transformed by a vector according to (7) above.

The invention still further provides (9) a protein encoded by a gene according to any one of (1) to (6) above.

The invention still further provides (10) a method of producing a protein having activity of transferring a sugar to the chalcone 4'-position, characterized by culturing or growing host cells according to (7) above and obtaining the protein from the host cells.

The invention still further provides (11) a plant having a gene according to any one of (1) to (6) above introduced therein or a progeny of the plant having the same properties as the plant, or tissue of such a plant.

The invention still further provides (12) a flower cut from a plant according to (11) above.

The invention still further provides (13) a method for transferring a sugar to the chalcone 4'-position using a gene according to any one of (1) to (6) above.

The invention still further provides (14) a plant having modified flower color obtained by introducing and expressing a gene according to any one of (1) to (6) above into the plant, or a progeny of the plant having the same properties as the plant.

The invention still further provides (15) a plant according to (14) above characterized in that the flower color has a yellow tint.

The invention still further provides (16) a method of introducing and expressing a gene according to any one of (1) to (6) above together with a gene coding for aureusidin synthase in a plant to alter the flower color to yellow.

The invention still further provides (17) a method of introducing and expressing a gene according to any one of (1) to (6) above together with a gene coding for aureusidin synthase in a plant, and also inhibiting expression of a flavonoid synthesis pathway gene in the

host, to alter the flower color to yellow.

The invention still further provides (18) a method of introducing and expressing a gene according to any one of (1) to (6) above together with a gene coding for aureusidin synthase in a plant, and also inhibiting expression of the dihydroflavonol reductase gene in the host, to alter the flower color to yellow.

The invention still further provides (19) a method of introducing and expressing a gene according to any one of (1) to (6) above together with a gene coding for aureusidin synthase in a plant, and also inhibiting expression of the flavanone 3-hydroxylase gene in the host, to alter the flower color to yellow.

Brief Explanation of the Drawings

Fig. 1 shows the typical flavonoid synthesis pathway in a plant. The details of the metabolic pathway differ for different plant species, depending on the presence or absence of the enzyme genes shown. For example, yellow *Antirrhinum majus* petals have both the anthocyanin synthesis pathway and pathway that leads to aurone synthesis, whereas torenia which is a plant belonging to the family *Scrophulariaceae* lacks the 4'CGT and AS genes and therefore cannot synthesize aurones. The abbreviations used in the drawing are explained below. CHS = chalcone synthase; CHI = chalcone isomerase; F3H = flavanone 3-hydroxylase; DFR = dihydroflavonol 4-reductase; ANS = anthocyanidin synthase; 3GT = UDP-glucose:anthocyanidin 3-glucosylase; FLS = flavonol synthase; FNS = flavone synthase; F3'H = flavonoid 3'-hydroxylase; F3',5'H = flavonoid 3',5'-hydroxylase; 2'CGT = UDP-glucose:4,2',4',6'-tetrahydroxychalcone 2'-glucosyltransferase; 4'CGT = UDP-glucose:4,2',4',6'-tetrahydroxychalcone 4'-glucosyltransferase; AS = aureusidin synthase.

Fig. 2 is a continuation of Fig. 1.

Fig. 3 shows the results of Southern hybridization

of a torenia transformant. Genomic DNA was extracted from pSFL201, pSFL307 and pSFL308 introduced torenia leaves (variety Summerwave Blue) transformants, and after KpnI cleavage, was subjected to Southern hybridization using the 4'CGT gene as the probe. The numerals at the top of each lane indicate the lineage numbers of the transgenic constructs and transformants. SWB represents Summerwave Blue used as the host. M1 and M2 are DIG-labeled size markers (Roche), and the size of each band is shown at left.

Best Mode for Carrying Out the Invention

As an example of a gene according to the invention there may be mentioned one coding for the amino acid sequence listed as SEQ ID NO: 2. However, it is well known that a protein having a modified amino acid sequence with one or a plurality of amino acids added, deleted and/or substituted with other amino acids can retain the enzyme activity of the original protein. The present invention therefore also includes any protein having the amino acid sequence listed as SEQ ID NO: 2 with a modification of one or a plurality of amino acids that are added, deleted and/or substituted with other amino acids, so long as it is a protein that retains 4'CGT activity, as well as a gene coding for such a protein. Here, "a plurality" means 2-30, and preferably 2-9.

The present invention also relates to a gene that hybridizes to DNA having the nucleotide sequence listed as SEQ ID NO: 1 under the relatively mild conditions of 5xSSC, 50°C, and that codes for a protein having 4'CGT activity. Also included within the technical scope of the invention is any gene that hybridizes to DNA having the nucleotide sequence listed as SEQ ID NO: 1 under stringent conditions, and that codes for a protein having 4'CGT activity. Here, the "stringent conditions" may be, for example, 2xSSC, 65°C, but the hybridization conditions

are not limited thereto because they may differ depending on the length and base composition of the DNA used as probe.

As genes selected by such hybridization there may be mentioned naturally occurring genes such as plant-derived genes, preferably *Scrophulariaceae*-derived genes, and more preferably *Antirrhinum majus*, *Linaria bipartita* and *Linaria japonica*-derived genes, although there is no limitation to plants. That is, the 4'CGT gene of the invention is not limited to the 4'CGT gene from *Antirrhinum majus*, *Linaria bipartita* or *Linaria japonica*, and the 4'CGT gene from other biological species that contain chalcone 4'-glucosides may be used to cultivate yellow-colored flowers. Synthetic DNA containing the 4'CGT gene may also be used in the same manner as plant-derived genes.

The gene selected by hybridization may be cDNA or genomic DNA.

As shown in the examples, a gene having homology with the conserved region of GT may be obtained by screening a cDNA library prepared from, for example, *Antirrhinum majus* or *Linaria bipartita* flower petals. DNA coding for GT having an amino acid sequence which is a modification of the amino acid sequence listed as SEQ ID NO: 2 may be synthesized by a publicly known site specific mutagenesis or PCR method using DNA having the nucleotide sequence listed as SEQ ID NO: 1. For example, a DNA fragment for amino acid sequence modification may be obtained by restriction endonuclease treatment of cDNA or genomic DNA and used as template for site specific mutagenesis or PCR using a primer corresponding to the desired modified amino acid sequence, in order to obtain a DNA fragment corresponding to the desired modified amino acid sequence. The modification-introduced DNA fragment may then be ligated to a DNA fragment coding for another portion of the enzyme of interest. Such DNA may also be chemically synthesized.

Alternatively, in order to obtain DNA coding for a protein comprising a shortened amino acid sequence, DNA coding for an amino acid sequence longer than the target amino acid sequence, for example, the full-length amino acid sequence, may be cleaved with a desired restriction endonuclease, and if the resulting DNA fragment does not code for the full-length target amino acid sequence, it may be ligated with a synthesized DNA fragment corresponding to the lacking portion of the amino acid sequence.

The GT gene obtained in this manner is expressed using a gene expression system in *E. coli* or yeast, and measurement of 4'CGT activity in the *E. coli* or yeast extract can confirm that the obtained GT gene codes for a protein exhibiting 4'CGT activity. The 4'CGT activity may be measured in the manner described in Example 3, for example, by adsorbing a chalcone serving as the substrate for 4'CGT onto a reverse-phase resin and then reacting the reverse-phase resin with extract of *E. coli* or yeast that has been transformed with the GT gene, and analyzing the produced chalcone 4'-glucoside by high performance liquid chromatography (HPLC).

The obtained 4'CGT gene may be expressed in suitable host cells to obtain 4'CGT protein as the gene product. Alternatively, an antibody for a protein or peptide having all or a portion of the amino acid sequence listed as SEQ ID NO: 2 may be used to obtain the 4'CGT gene of another organism by expression cloning.

The present invention further relates to a recombinant vector, especially an expression vector, containing the 4'CGT gene, and to host cells transformed using the vector. The host cells used may be prokaryotic or eukaryotic. As examples of prokaryotic cells there may be used publicly known host cells including bacteria belonging to the genus *Escherichia* such as *Escherichia coli*, or microorganisms belonging to the genus *Bacillus* such as *Bacillus subtilis*.

As examples of eukaryotic cells there may be used eukaryotic microorganisms, and preferably yeast or filamentous fungi. As examples of yeast there may be mentioned *Saccharomyces* yeast such as *Saccharomyces cerevisiae*, and as examples of filamentous fungi there may be mentioned *Aspergillus* microbes such as *Aspergillus oryzae* and *Aspergillus niger*, or *Penicillium* microbes. Animal or plant cells may also be used as host cells, among which mouse, hamster, monkey or human cell lines may be used as the host cells. Insect cells such as silkworm cells, or actual silkworm adults, may be used as hosts.

The expression vector of the invention includes expression regulating regions such as a promoter and terminator dependent on the host species, as well as a replication origin. As promoters for the expression vector in bacteria, and particular in *E. coli*, there may be used conventional publicly known promoters such as trc promoter, tac promoter and lac promoter. As yeast promoters there may be used, for example, glyceraldehyde 3-phosphate dehydrogenase promoter, PH05 promoter or the like and as filamentous fungi promoters there may be used, for example, amylase or trpC promoters, but there is no limitation to these promoters. As animal cell promoters there may be used viral promoters, such as SV40 early promoter or SV40 late promoter. The expression vector may be constructed by ordinary methods using restriction endonucleases, ligases and the like. Transformation of the host cells with the expression vector may also be carried out by conventional publicly known procedures.

Construction of a plant expression vector may be accomplished, for example using a binary vector such as pBI121 when using *Agrobacterium*, or using an *E. coli* vector such as pUC19 when using a particle gun. The plant cells transformed with the plant expression vector may then be selected using a marker gene such as an

antibiotic resistance gene and regenerated under conditions employing a suitable plant hormone, to obtain transformed plant individuals having the 4'CGT gene introduced therein. Cultivation of the transformed plants to blooming can yield plants exhibiting the modified flower color.

The host cells transformed by the expression vector or the transformed plants may be cultured or cultivated, and the target 4'CGT protein recovered and purified from the culture, etc. by ordinary methods such as, for example, filtration, centrifugal separation, cell disruption, gel filtration chromatography, ion-exchange chromatography and the like.

The present invention is not limited only to the 4'CGT gene of *Antirrhinum majus* or *Linaria bipartita*, but may also be applied for modification of flower color by 4'CGT whether the source of the 4'CGT or 4'CGT gene is a plant, animal, microorganism or synthesized organism. The present invention relates to the use of the 4'CGT gene, and the scope of the invention encompasses plants with modified flower color by introduction and expression of the 4'CGT gene in plants, as well as their progeny, vegitatively propagated products of the foregoing and tissues of such plants, where the tissues may include cut flowers. Moreover, the scope of the present invention further encompasses plants with modified flower color by introduction and expression in plants not only of the 4'CGT gene but even AS genes other than the 4'CGT gene, and by inhibited expression of the endogenous flavonoid synthesis pathway genes of the host, as well as their progeny, vegitatively propagated products of the foregoing and tissues of such plants, where the tissues may include cut flowers.

With the current level of technology it is possible to introduce genes into plants and express those genes either constitutively or in a tissue-specific manner, or to inhibit expression of target genes by antisense,

cosuppression or RNAi methods. As examples of plants to be transformed there may be mentioned rose, chrysanthemum, carnation, snapdragon, cyclamen, morning glory, begonia, impatiens, geranium, orchid, bluebell, freesia, gerbera, gladiolus, gypsophila, kalanchoe, lily, pelargonium, geranium, petunia, torenia, tulip, rice, forsythia, begonia, barley, wheat, rapeseed, potato, tomato, poplar, banana, eucalyptus, sweet potato, soybean, alfalfa, lupin, corn and cauliflower, although there is no limitation to these.

Examples

The present invention will now be explained in greater detail by the following examples. The molecular biological methods employed were based on the description in WO96/25500 or Molecular Cloning (Sambrook et al. Cold Spring Harbour Laboratory Press, 1989), unless otherwise specified.

Example 1. Construction of yellow *Antirrhinum majus* flower petal cDNA library

A cDNA library was constructed as described in the literature (Science 290, 1163 (2000)), from 5 g of fresh flower petals of the yellow *Antirrhinum majus* variety Butterfly Yellow. The obtained library consisted of 1.6 x 10^5 plaque forming units.

Example 2. 4'CGT gene screening 1

Previously disclosed GT amino acid sequences were compared, and nucleotide sequences corresponding to the conserved regions of the amino acid sequences were amplified and used as probes for screening of the *Antirrhinum majus* cDNA library described in Example 1.

Five GTs were used as the probes, namely the sequences for morning glory-derived UDP-glucose:anthocyanidin 3-glucoside glucosyltransferase (3GGT) (Japanese Unexamined Patent Publication No. 2003-289884), petunia-derived 3GT (Plant Mol. Biol. 48, 401, (2002)), verbena-derived 5GT (J. Biol. Chem. 274, 7405

(1999)), Scutellaria baicalensis-derived GT (SBGT, Planta 210,1006 (2000)) and gentian-derived UDP-glucose:anthocyanin 3'-glucosyltransferase (3'GT) (Plant Physiol. 132,1652, (2003)). One oligonucleotide pair was synthesized for each GT to allow amplification of the conserved region sequence. The oligonucleotide sequences are listed as SEQ ID NO: 3-12.

Morning glory 3GGT

SEQ ID NO: 3: 5'-GAA ATG GTC GGA TTG GCT GGG-3'

SEQ ID NO: 4: 5'-ACC TCC ACC CCA ACT TTC AGG-3'

Petunia 3GT

SEQ ID NO: 5: 5'-GAT GCA TAA TTT GGC TAG AAA AGC-3'

SEQ ID NO: 6: 5'-CCA ATT TGC CAA ACA CTT TCC-3'

Verbena 5GT

SEQ ID NO: 7: 5'-TGC CTC GAA TGG TTG AGC ACG-3'

SEQ ID NO: 8: 5'-CTC TCA CTC TCA CAC CCG-3'

Baikal skullcap GT

SEQ ID NO: 9: 5'-CAC GAA TGC TTA GCA TGG CTC-3'

SEQ ID NO: 10: 5'-CTT ATT GCC CAC TGA AAC CCC-3'

Gentian 3'GT

SEQ ID NO: 11: 5'-TGT CTG AAT TGG CTT GAT TCC-3'

SEQ ID NO: 12: 5'-AAC CCA CAG AAA CCC CTG TTC-3'

The probes were labeled using a non-radioisotope DIG-nucleic acid detection system (Roche Diagnostic Corp.), with PCR under the conditions recommended by the manufacturer. Plasmids containing 1 ng of each cDNA were used as template and 100 ng of oligonucleotide specific to each gene was used as primer, for 25 cycles of PCR where each cycle consisted of reaction at 95°C for 1 minute, 55°C for 1 minute and 72°C for 2 minutes. An equivolume mixture of the PCR amplification product for each gene was used as the hybridization probe for screening of the *Antirrhinum majus* cDNA library of Example 1.

The hybridization was carried out overnight at 37°C in 5XSSC containing 30% formaldehyde, 1% SDS, and the

filter was rinsed for 30 minutes at 55°C using 5XSSC and 1% SDS. The positive signal by screening was detected using a non-radioisotope DIG-nucleic acid detection system (Roche Diagnostic Corp.), according to the method recommended by the manufacturer. Approximately 300,000 plaques were screened, leading to selection of 10 clones possessing full-length glucosyltransferase genes, which were designated as pSPB264, 1621, 1620, 1622, 1610, 1609, 1617, 1615, 660 and 658. A DNA Sequencer model 3100 (Applied Biosystems) was used to determine their cDNA sequences by the primer walking method with synthetic oligonucleotide primers. The nucleotide sequences of the amino acid-coding regions of the cDNA are listed as SEQ ID NO: 14-23.

Example 3. Assay of chalcone GT activity using *E. coli*

3-1 Construction of *E. coli* expression vector and expression of GT in *E. coli*

The activity of GT encoded by the ten cDNAs obtained in Example 2 was analyzed using an *E. coli* expression system. First, *E. coli* expression constructs were prepared for each cDNA. An NcoI site was introduced by PCR so as to be situated on the nucleotide sequence ATG believed to be the start codon for each cDNA, and the region from the start methionine to the end codon was linked to the NcoI and KpnI or NcoI and EcoRV sites of the *E. coli* expression vector pQE61 (QIAGEN).

The PCR solution (25 μl) for introduction of the NcoI site situated on the start methionine comprised each GT cDNA as template, 0.2 pmol/μl each of primer containing the NcoI recognition site situated on the start methionine position and primer from the 3' end to the 5' end near the stop codon, 1x ExTaq buffer (Takara), 0.2mM dNTPs and 1.25 U ExTaq polymerase. Reaction was conducted at 94°C for 5 minutes, followed by 28 cycles of reaction at 94°C for 1 minute, 55°C for 1 minute and 72°C

for 2 minutes, and then by final treatment at 72°C for 5 minutes. The obtained PCR product was subcloned in pCR2.1 TOPO vector (INVITROGEN) according to the method recommended by the manufacturer. The DNA sequences of the amplified DNA fragments were analyzed, and after confirming lack of PCR-induced error, they were introduced into *E. coli* expression vector pQE61 (QIAGEN).

For example, for cDNA encoded in pSPB1617 (SEQ ID NO: 20), the two different primers 1617 BamHINcoI-FW (SEQ ID NO: 24) and the 1617 XhoIKpnI-RV (SEQ ID NO: 25) listed in the Sequence Listing were used for PCR, for introduction of the NcoI site lying at the start methionine position and the KpnI recognition sequence at the 3'-end of the end codon. The amplified DNA fragments were subcloned in pCR2.1 TOPO vector. After confirming lack of PCR-induced error, the DNA fragment cut out with NcoI and KpnI was linked to the NcoI and KpnI sites of pQE61, to obtain pSPB1642 as a pSPB1617cDNA *E. coli* expression vector. Ten different GT cDNA *E. coli* expression vectors were constructed in the same manner.

1617BamHINcoI-FW

SEQ ID NO: 24: 5'-GGG GGA TCC ATG GCT AGT GAG AGC CAA ATA-3'

1617XhoIKpnI-RW

SEQ ID NO: 25: 5'-CCC CTC GAG GGT ACC TCA CAA AAC ATT ATT CAC GAC-3'

Each expression vector was introduced into *E. coli* JM109 (TOYOBO) and pre-cultured overnight at 37°C in LB medium containing ampicillin at a final concentration of 20 μg/ml. After adding 1 ml of pre-culturing solution to 50 μg/ml of ampicillin and 50 ml of M9 medium containing 0.5% casamino acid and culturing to A600 = 0.6-1.0, IPTG (Isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 0.1 mM and shake culturing was carried out overnight at 27°C, and was followed by centrifugation at 3000 rpm for 10 minutes at 4°C to

recover the cells. The cells were suspended in 10 ml of buffer solution (30 mM Tris-HCl pH 7.5, 30 mM NaCl), and after disruption of the *E. coli* with an ultrasonic treatment with a SONIFIER 250 (BRANSON), it was centrifuged at 15,000 rpm, 10 min, 4°C and the obtained supernatant was used as the crude enzyme solution for the following activity assay.

3-2 Enzyme activity assay

After loading THC (500 μg/ml ethanol solution) into 1 ml of the reverse-phase resin TOYOPEARL HW-40F (TOSOH) equilibrated with H_2O, while diluting with H_2O, it was rinsed with water to obtain THC substrate fixed on the resin. To 100 μl of the resin-fixed THC there were added 200 μl of the crude enzyme solution obtained in 3-1 and 10 μl of 5 mM UDP-glucose, and reaction was conducted at 30°C for 1 hour. After removing the supernatant by centrifugation, the precipitated resin was rinsed with water and suspended in 300 μl of 50% acetonitrile containing 0.1% TFA (Trifluoroacetic acid), and the flavonoids were freed from the resin by ultrasonic treatment. Upon centrifugation at 15,000 rpm, 5 min, 4°C, the insoluble portion of the obtained supernatant was removed with a filter (pore size: 0.45 mm, 4 mm Millex-LH, Millipore), and the supernatant was analyzed by high-performance liquid chromatography (HPLC). The analysis conditions for chalcones and their glucosides were as follows.

The column used was a Develosil C-30-UG-5 (4.5 mmϕ x 150 mm, Nomura Chemical Co., Ltd.), with a mobile phase of H_2O containing 0.1% TFA as solution A and 90% acetonitrile containing 0.1% TFA as solution B, and after elution for 10 minutes with a linear concentration gradient from 20% solution B to 70% solution B, it was held for 5 minutes with 70% solution B. The flow rate was 0.6 ml/min, and detection was carried out based on absorbance at 360 nm and the absorption spectrum from

250-400 nm using an SPD-M6A PDA detector (Shimadzu Corp.). Elution of THC at a retention time of 10.7 min, and 2'-glucoside and 4'-glucoside at 8.5 min under these conditions was confirmed using samples of THC and THC 2'- and 4'-glucosides.

Upon reaction of the pSPB1642-expressing *E. coli* extract, a new product in addition to the THC substrate was detected, eluting at 8.5 min. Since these were not detected with reaction using crude extract prepared in the same manner from *E. coli* expressing only pQE61 vector or boiled crude extract of *E. coli* expressing pSPB1642, it was concluded that the product was produced by GT expressed from pSPB1642. The structure of the product was also examined by ^{1}H-NMR analysis. The analysis was performed using a JNM-EX400 (JEOL), under conditions as described in the literature (Plant Physiology 132, 1652 (2003)). As a result, the THC glucoside produced by the expression product of pSPB1642 was demonstrated to be THC 2'-glucoside. Thus, it was concluded that the cDNA expressed by pSPB1642, i.e. pSPB1617 cDNA, codes for a protein with 2'CGT activity.

Example 4. 4'CGT gene screening 2

The approximately 300,000 clones of the yellow *Antirrhinum majus* flower petal cDNA library were re-screened using the full-length pSPB1617 cDNA as the probe. The probe labeling by PCR was conducted in the same manner as described in Example 2, using 1617-F (SEQ ID NO: 26) and 1617-R (SEQ ID NO: 27) primers. The screening and nucleotide sequence analysis methods were also the same as in Example 2.

1617-F

SEQ ID NO: 26: 5'-ATG GGA GAA GAA TAC AAG AAA ACA-3'

1617-R

SEQ ID NO: 27: 5'-TAA AAT TTG GTA GTT AAA CCG ATG TA-3'

As a result, five novel GT genes were obtained: pSPB1721, 1724, 1723, 1719 and 1725. The respective

sequences are listed in the Sequence Listing (SEQ ID NO: 28-31 and 1).

Among these, pSPB1725 cDNA contained a 1374 bp translation region (excluding the stop codon) coding for a 50.8 kDa molecular weight protein composed of 457 amino acids, with an isoelectric point of 6.82. The amino acid sequence (SEQ ID NO: 2) encoded by the pSPB1725 cDNA was compared with the previously reported GT amino acid sequence, and only showed 14% homology with Livingstone Daisy GT (Plant J. 19, 509 (1999)), 18% homology with perilla 5GT, 18% homology with perilla 3GT, 23% homology with gentian 3'GT and 31% homology with the amino acid sequence of the protein encoded by pSPB1617 used as the probe. The software used for homology analysis was ClustalW included with MacVector ver.6.5.3 (Oxford Molecule), with parameters of Matrix Blosum 30, ketuple: 1, Gap penalty: 3, Topdiagonals: 5, Windows Size: 5.

Example 5. Expression of obtained cDNA in *E. coli*

5-1. Expression vector construction

The five cDNAs obtained in Example 4 were examined by assaying the enzyme activity of the protein encoded by each cDNA using an *E. coli* expression system. The expression vector construction and expression method and activity assay method were the same as in Example 3. For pSPB1725, for example, PCR reaction was conducted using the following two primers 1725-NcoI (SEQ ID NO: 32) and 1725-KpnI (SEQ ID NO: 33) for introduction of the NcoI recognition sequence at the 5'-end of the start codon.

1725-NcoI

SEQ ID NO: 32: 5'-CCC ATG GGA GAA GAA TAC AAG AAA-3'

1725-KpnI

SEQ ID NO: 33: 5'-GGT ACC TAT AAA ATT TGG TAG TTA AA-3'

The PCR solution (25 µl) consisted of 10 ng pSP1725 DNA, 1x ExTaq buffer (Takara), 0.2 mM dNTPs, 0.2 pmol/µl each of the 1725-NcoI and 1725-KpnI primers and 1.25 U ExTaq polymerase.

Reaction was conducted at 94°C for 5 minutes, followed by 28 cycles of reaction at 94°C for 1 minute, 55°C for 1 minute and 72°C for 2 minutes, and then by final treatment at 72°C for 7 minutes. The obtained PCR product was subcloned in pCR2.1 TOPO vector (INVITROGEN) according to the method recommended by the manufacturer. Upon confirming the nucleotide sequence of the amplification product, the approximately 1.4 kb fragment cut out from pCR2.1 TOPO vector by NcoI and KpnI treatment was linked to the NcoI and KpnI sites of pQE61 (QIAGEN) to obtain the *E. coli* expression vector pSPB1768. This was then introduced into *E. coli* JM109 (TOYOBO). *E. coli* expression vectors were also constructed for the other four cDNAs in the same manner using pQE61, and were introduced into JM109.

5-2 Recombinant protein expression in *E. coli* and GT activity assay

The *E. coli* transformants obtained in Example 5-1 were cultured under the same conditions as Example 3 and used for activity assay of the proteins encoded by each cDNA. As a result, a peak attributed to THC glucoside was detected in the reaction product between THC and the crude enzyme solution of pSPB1768-containing *E. coli*. In order to identify the THC glucoside in greater detail, HPLC analysis was again performed under conditions for separation of THC 2'-glucoside and THC 4'-glucoside, as follows.

The column used was a YMC-ODS-A312 (6 mmϕ x 150 mm, YMC Corp.), and with a mobile phase of H_2O containing 2% acetic acid as solution A and methanol as solution B, for elution for 15 minutes with a linear concentration gradient from 15% solution B to 40% solution B and holding for 5 minutes with 40% solution B, followed by elution for 10 minutes with a linear concentration gradient from 40% solution B to 62% solution B and holding for 2 minutes with 62% solution B. The flow rate

was 1.0 ml/min. Detection was carried out based on absorbance at 360 nm and the absorption spectrum from 250-400 nm using an SPD-M6A PDA detector (Shimadzu Corp.).

Under these conditions, THC elutes at a retention time of 26.7 min, THC 2'-glucoside elutes at 19.8 min and 4'-glucoside elutes at 20.6 min. The THC glucoside found in the pSPB1768-expressing *E. coli* extract and THC reaction solution was believed to be 4'-glucoside because it eluted at 20.6 min by analysis under these conditions. Since it was not detected upon reaction of crude extract prepared in the same manner from *E. coli* expressing only pQE61 vector, it was presumably the product of GT encoded by pSPB1725. These results confirmed that GT encoded by pSPB1725 cDNA exhibits activity of transferring glucose to the 4'-hydroxyl group of THC.

A new peak for elution at 15.5 min was also detected in addition to THC 4'-glucoside in the reaction solution. This substance exhibited the absorption spectrum of naringenin, and matched the retention time for a naringenin 7-glucoside standard. Thus, the product that eluted at 15.5 min was believed to be naringenin 7-glucoside produced by post-glucosylation ring closure of THC 4'-glucoside produced by 4'CGT encoded in pSPB1725, or naringenin 7-glucoside produced by the action of 4'CGT on naringenin produced by ring closure of THC.

Example 6. Expression analysis of 4'CGT gene in *Antirrhinum majus* flower petals

The expression profile of the 4'CGT gene encoded by pSPB1725 in yellow *Antirrhinum majus* flower petals was analyzed by RT-PCR. Flower petals of yellow *Antirrhinum majus* (Butterfly Yellow variety) which accumulate aurones were separated at 5 growth stages. The five stages are, in order of youth, stages 1 (budding petal length of ≤1 cm), 2 (budding petal length of 1.0-1.5 cm), 3 (budding petal length of 1.5-2.0 cm), 4 (petal length of 2.0-2.5 cm, just prior to blooming) and 5 (petal length of ≥2.5

cm, bloomed petals), where stage 5 corresponds to the mature flower petal.

The RNA was extracted from 1 g of separated flower petals, using an RNeasy Plant Mini Kit (QIAGEN). Reverse transcription reaction was conducted using 1 μg of obtained RNA as template to obtain cDNA. A SuperScript First-Strand Synthesis System for RT-PCR (GIBCO BRL) was used for the cDNA synthesis, and the synthesis conditions were according to the conditions recommended by the system manufacturer. PCR was performed using the cDNA obtained at each stage as template and the 1725-NcoI (SEQ ID NO: 32) and 1725-KpnI (SEQ ID NO: 33) described in Example 5 as primers. For comparison between 4'CGT gene expression and endogenous gene expression, the *Antirrhinum majus* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene (SEQ ID NO: 34) (Nature 339, 46 (1989)) was used as the internal standard gene, and primers AmGAPDH-F (SEQ ID NO: 35) and AmGAPDH-R (SEQ ID NO: 36) were synthesized for amplification of this gene. Also, primers AmAS-F (SEQ ID NO: 37) and AmAS-R (SEQ ID NO: 38) were synthesized for amplification of the *Antirrhinum majus* AS gene, as a comparison gene.

AmGAPDH-F

SEQ ID NO: 35:5'-TGT TGC TGT TAA CGA TCC AT-3'

AmGAPDH-R

SEQ ID NO: 36:5'-AGC TCT TCC ACC TCT CCA-3'

AmAS-F

SEQ ID NO: 37:5'-ATG TTC AAA AAT CCT AAT ATC CGC-3'

AmAS-R

SEQ ID NO: 38:5'-TTA GCC ATC AAG CTC AAT CTT GAC A-3'

The PCR conditions were 12 cycles of 94°C for 1 minute, 55°C for 1 minute and 72°C for 2 minutes, with the same reaction composition as in Example 3. The PCR product was separated by 1% agarose gel electrophoresis, and then blotted on a Hybond-N nylon membrane (Amersham)

according to ordinary protocol for detection of amplification product by hybridization. The hybridization method employed the previously mentioned non-radioisotope DIG-nucleic acid detection system, DIG DNA labeling and detection kit, and was carried out according to the method recommended by the manufacturer. The probes used were cDNA for 4'CGT encoded in *Antirrhinum majus* AS, GAPDH and pSPB1725, and DIG labeling was accomplished using primers specific for each gene (SEQ ID NO: 32, 33, 35-38) in the same manner as Example 2.

As a result, both the 4'CGT and AS genes reached expression peaks at stage 4, thus exhibiting similar expression patterns with time. The expression patterns for both genes also did not appear inconsistent with the accumulation patterns for chalcone 4'-glucosides and aurones in yellow *Antirrhinum majus* flower petals (Plant Sci. 160, 229 (2001)).

These results suggest that expression of the 4'CGT gene encoded in pSPB1725 is under the same control as the AS gene in *Antirrhinum majus* flower petals, and that both are involved in the same synthesis pathway, namely the synthesis pathway of aurone.

Example 7. Co-expression of 4'CGT and AS in plant

7-1 Construction of 4'CGT expression cassette

Plasmid pBE2113-GUS (Plant Cell Physiol. 37, 45 (1996)) was digested with SnaBI and religated to remove the omega sequence, and the obtained plasmid was designated as pUE6. Separately, plasmid pUCAP (van Engelen et al. Transgenic Research 4, 288-290, 1995) was digested with AscI and the ends blunted, after which PacI linker was inserted to prepare a plasmid designated as pUCPP. A fragment of pUE6 from the $El_2$35S promoter to the NOS terminator was inserted at the HindIII and EcoRI sites of pUCPP to obtain plasmid pSPB540. The GUS gene portion of pSPB540 was replaced with the 4'CGT cDNA fragment cut out from pSPB1725, and the obtained plasmid

was designated as pSFL203. That is, pSFL203 comprises pUCPP as the vector and has the 4'CGT expression cassette controlled by $El_2$35S promoter and NOS terminator.

7-2 Construction of AS expression cassette

Antirrhinum majus-derived AS cDNA (Science 290, 1163, (2000)) was inserted at the EcoRI and XhoI sites of pBluescript II SK-vector (Stratagene) to obtain a plasmid designated as pSPB251. An AS expression construct having the MacI promoter, the AS cDNA fragment cut out from pSPB251 and the MAS terminator linked in plasmid pBINPLUS (van Engelen et al. Transgenic Research 4, 288-290, 1995) was designated as pSPB1624.

7-3 Creation of 4'CGT and AS co-expression construct

The pSFL203 of 7-1 was cleaved with PacI and the chalcone glucosylating enzyme gene expression cassette was cut out and inserted at the PacI site of plasmid pSPB1624 from 7-2. The obtained construct was designated as pSFL201. Thus, pSFL201 is designed so as to constitutively express the 4'CGT and AS genes when introduced into plant cells.

Example 8 Co-expression of 4'CGT and AS and inhibition of torenia DFR in plants

8-1 Construction of torenia-derived DFR gene expression-inhibiting cassette

The cDNA for torenia dihydroflavonol reductase (DFR) was obtained in the manner described in the literature (Plant Science 153, 33, 2000). The torenia DFR cDNA was linked with the vector pBluescriptII SK- to obtain a plasmid designated as pTDF10. This was used as a template for PCR in the manner described in Example 3, using M13 reverse primer (SEQ ID NO: 39) from the vector sequence and the primer ThDFR-NcoI (SEQ ID NO: 40) having an NcoI recognition site introduced therein by nucleotide substitution in the torenia DFR cDNA sequence. The approximately 0.75 kb fragment that was obtained was cloned in pCR2.1-TOPO (Invitrogen), and after confirming the nucleotide sequence, the 0.75 kb torenia DFR cDNA

sequence was cut out with SacI and NcoI.

Also, plasmid pTDF10 was cleaved with BamHI and NcoI, and a fragment containing 1.1 kb was recovered from the 5' end of torenia DFR cDNA. Separately, plasmid pUCAP described in 7-1 was digested with PacI and its ends blunted, after which an AscI linker was inserted to obtain a plasmid designated as pUCAA. A fragment including $El_2$35S promoter-GUS-NOS terminator cut out from pUE6 was inserted at the HindIII and EcoRI sites of pUCAA, and the obtained plasmid was designated as pSPB541. Plasmid pSPB541 was cleaved with BamHI and SacI, the GUS gene portion was removed, and there were inserted therein the 0.75 kb fragment and 1.1 kb fragment from torenia DFR cDNA, ligated in a direction linking the NcoI sites of both fragments. Plasmid pSFL314 obtained in this manner, when introduced into a plant, can transcribe double-stranded RNA from the torenia DFR cDNA sequence under the control of $El_2$35S promoter, thus inhibiting expression of the torenia DFR gene by RNAi.

M13 reverse primer

SEQ ID NO: 39: 5'-AACAGCTATGACCATG-3'

ThDFR-NcoI

SEQ ID NO: 40: 5'-GCTTTACCATGGAGTAATGAGCTT-3'

8-2 Co-expression of 4'CGT and AS and creation of construct for inhibition of torenia DFR gene expression

An XhoI linker was inserted upstream of the NOS terminator of pUE6 described in 7-1. The plasmid was digested with BamHI and XhoI to obtain a fragment comprising the $El_2$35S promoter-vector-NOS terminator, and this fragment was ligated with the AS cDNA fragment cut out from pSPB215 with BamHI and XhoI as described in 7-2 to obtain plasmid pSPB211. The AS expression cassette was cut out from pSPB211 with HindIII and EcoRI and inserted at the HindIII and EcoRI sites of pBINPLUS. The 4'CGT expression cassette obtained by PacI cleavage of pSFL203 described in 7-1 was inserted at the PacI site of the plasmid obtained in this manner, to obtain plasmid

pSFL304 having the 4'CGT and AS expression cassettes linked in tandem. Also, the torenia DFR double-stranded RNA transcription cassette described in 8-1 was inserted at the AscI site of pSFL304 to obtain plasmid pSFL307. That is, pSFL307 includes three cassettes, for 4'CGT and AS expression and torenia DFR gene expression inhibition.

Example 9 Co-expression of 4'CGT and AS and inhibition of torenia F3H gene expression in plant

9-1 Cloning of torenia F3H cDNA and construction of expression inhibition cassette for the gene

F3H cDNA obtained from perilla (Plant Mol Biol., 35, 915 (1997)) was used as a probe to obtain cDNA coding for the same enzyme in torenia. Specifically, a torenia cDNA library (Molecular Breeding, 6, 239, 2000) of approximately 200,000 phage was screened in the same manner as Example 2 to obtain the torenia F3H cDNA listed as SEQ ID NO: 41. The torenia F3H cDNA was linked with the vector pBluescriptII SK- to obtain a plasmid designated as pSPB266. This was used as template for PCR in the manner as Example 3, using M13 reverse primer (SEQ ID NO: 39) from the vector sequence and the primer ThF3H-SalI-1 (SEQ ID NO: 42) having a SalI recognition site inserted therein by nucleotide substitution in the torenia F3H cDNA sequence.

The approximately 0.9 kb fragment that was obtained was cloned in pCR2.1-TOPO (Invitrogen), and the nucleotide sequence was confirmed. In the same manner, the primer ThF3H-SalI-2 (SEQ ID NO: 43) having a SalI recognition site inserted therein by nucleotide substitution in the torenia F3H cDNA sequence and M13 reverse primer were used to prepare an approximately 0.75 kb DNA fragment which was cloned in pCR2.1-TOPO, and the nucleotide sequence was confirmed. Plasmid pSPB541 described in Example 8-1 was cleaved with BamHI and SacI, the GUS gene portion was removed, and there were inserted therein a 0.9 kb fragment cut out from pCR2.1-TOPO by cleavage with BamHI and SalI and a 0.7 kb fragment cut

out from pCR2.1-TOPO by cleavage with SacI and SalI, ligated in a direction linking the SalI sites of both fragments. Plasmid pSFL313 obtained in this manner, when introduced into a plant, transcribes double-stranded RNA derived from the torenia F3H cDNA sequence under the control of $El_2$35S promoter, thus inhibiting expression of the torenia F3H gene by RNAi.

ThF3H-SalI-1

SEQ ID NO: 42: 5'-TTCTCTGTCGACGCCCATTGCC-3'

ThF3H-SalI-2

SEQ ID NO: 43: 5'-CGCCGTGTCGACTCGCTTGAAG-3'

9-2 Co-expression of 4'CGT and AS and creation of construct for inhibition of torenia F3H gene expression

The torenia F3H RNAi cassette was cut out from the pSFL313 described in 9-1 by cleavage with AscI and inserted at the AscI site of pSFL304 described in Example 8-2 to obtain plasmid pSFL308. That is, pSFL308 includes three cassettes, for 4'CGT and AS expression and torenia F3H gene expression inhibition.

Example 10. Gene expression and flower color analysis in plant

Plasmids pSFL201, pSFL307 and pSFL308 described in Example 7-9 were introduced into torenia (variety: Summerwave Blue (Suntory Flowers Ltd.)) by a publicly known method. The transformation method was based on the method described in Mol. Breeding. 6, 239, (2000). Individuals exhibiting selection marker resistance were selected out and their flower colors observed. Among the pSFL201-introduced varieties, 22 of the 35 transformant lines exhibited altered flower color with respect to the host, displaying a yellowish blue or yellowish gray color.

However, no pure yellow color was found. Among the pSFL307-introduced varieties, 19 of the 36 transformant lines exhibited altered flower color with respect to the host. Also, 6 of the 19 lines with altered flower color displayed almost pure yellow flower colors virtually

without any tint of the original blue color of the host. Among the pSFL308-introduced varieties, 24 of the 39 transformant lines exhibited altered flower color with respect to the host. Also, 17 of the 24 lines with altered flower color displayed almost pure yellow flower colors virtually without any tint of the original blue color of the host.

Pigment analysis was performed on the lines with relatively notable flower color alteration.

The flower petals of the host variety and the transformants were immersed with 50% acetonitrile containing 0.1% trifluoroacetic acid (TFA) for extraction of the flavonoids, and then the aureusidin 6-glucosides and anthocyanidins were analyzed by high performance liquid chromatography (HPLC). For anthocyanidin analysis, the flavonoids extracted from the flower petals were dissolved in 6N HCl and hydrolyzed by holding in boiling water for 20 minutes, after which the flavonoids were re-extracted with amyl alcohol and supplied for analysis. The HPLC conditions were as follows.

First, detection of AU 6-glucoside was performed using a SHIM-PACK FC-ODS column (50 x 4.6 mm, Shimadzu Corp.), and for the reverse phase, H_2O containing 0.05% TFA was used as solution A and acetonitrile containing 0.05% TFA was used as solution B. Elution for 3 minutes with a linear concentration gradient from 10%-23% solution B was followed by 17 minutes with 23% solution B, and then elution for 2 minutes with a linear concentration gradient from 23%-80% solution B was followed by 3 minutes with 80% solution B. Elution was then performed for 2 minutes with a concentration gradient from 80%-10% solution B. The flow rate was 0.8 ml/min. Detection was carried out based on absorbance at 360, 400 nm and the absorption spectrum from 250-500 nm using an SPD-M10AVP PDA detector (Shimadzu Corp.). Under these conditions, THC 4'-glucoside and AU 6-glucoside samples elute at retention times of 14.17 min and 6.19

min, respectively.

The column used for anthocyanidins was a YMC-ODS-A A312 (6 x 150 mm, YMC Corp.). For the reverse phase there was used a mixture of acetic acid, methanol and H_2O at 60:70:270, maintained for 11 minutes. Detection was carried out based on absorbance at 520 nm and the absorption spectrum from 400-600 nm using an SPD-M10AVP PDA detector (Shimadzu Corp.). Under these conditions, malvidin elutes at a retention time of 9.12 min.

As a result, with the pSFL201-introduced transformants, products matching the retention times and absorption spectra of THC 4'-glucoside and AU 6-glucoside were confirmed to be produced in the flower petals at 0.02% and 0.05% respectively (W/W as fresh flower petal weight). Because anthocyanidins native to the host were also present in the transformants, it may be assumed that the yellowish blue or gray colors observed in the transformants were due to the copresence of anthocyanidins such as malvidin with the THC 4'-glucoside and AU 6-glucoside.

On the other hand, with the pSFL307- and 308-introduced transformants, products matching only the retention times and absorption spectra of the aurone AU 6-glucoside were confirmed to be produced in the flower petals at 0.09% (W/W as fresh flower petal weight) in both cases. In the pSFL307- or pSFL308-introduced lines, it was confirmed that anthocyanidins native to the host were notably reduced to 10-50% of those anthocyanidins in the host flower petals.

Example 11. Confirmation of 4'CGT gene introduction by genomic Southern hybridization

Lines among the transformants obtained in Example 10 which had relatively large accumulation of THC 4'-glucoside and AU 6-glucoside based on the flower petal pigment analysis were selected (3 lines from each construct-introduced variety) and subjected to genomic hybridization. A Phytopure Plant DNA Extraction kit

(Amersham) was used for extraction of genomic DNA from approximately 1 g of transformant leaves, according to the method recommended by the manufacturer. Twenty μg of the obtained genomic DNA was cleaved with restriction endonuclease KpnI, and after separation by 0.7% agarose gel electrophoresis and transfer to a Hybond-N^+ nylon membrane according to ordinary protocols, a non-radioisotope DIG-nucleic acid detection system was used for hybridization.

The methods of DIG labeling, hybridization and detection of the 4'CGT gene probes were according to the methods recommended by the manufacturer, as in Example 6. The hybridization results are shown in Fig. 2. Judging from restriction enzyme maps for pSFL201, pSFL307 and pSFL308, it is possible to estimate the number of 4'CGT gene copies introduced into each transformant based on the number of bands detected by genomic Southern hybridization.

As one band was observed for all of the pSFL201-introduced lines, it is deduced that each has one copy of the introduced gene. Individuals of line Nos. 2 and 4 among the pSFL307-introduced lines had 1 copy, while individuals of line No. 13 exhibited 2 bands and therefore presumably have introduced 2 copies of 4'CGT cDNA. Since one band was observed for all of the pSFL308-introduced lines, it is deduced that each has one copy of the introduced gene.

Example 12. Introduced gene expression analysis by quantitative RT-PCR

An RNeasy Plant Mini Kit (QIAGEN) was used for extraction of total RNA from buds of the stock variety and each of the transformant lines, cDNA was synthesized from 1 μg of the obtained total RNA using a SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen). Using 1 μl of the obtained cDNA as template, expression of torenia DFR and F3H and transcription products of the

exogenous genes AS and 4'CGT was quantitatively analyzed with an ABI PRISM 7000 Sequence Detection System (Applied Biosystems). Using the manufacturer's recommended software "Primer Express", there were designed oligoprimers for specific amplification of each gene, and specific hybridizing TaqMan probes fluorescently labeled at both ends, and these were used for reaction. Quantitation of expression was carried out using the oligoprimers of SEQ ID NO: 54 and 55 and the TaqMan probe of SEQ ID NO: 56 for torenia DFR, the oligoprimers of SEQ ID NO: 57 and 58 and the TaqMan probe of SEQ ID NO: 59 for torenia F3H, the oligoprimers of SEQ ID NO: 60 and 61 and the TaqMan probe of SEQ ID NO: 62 for AS and the oligoprimers of SEQ ID NO: 63 and 64 and the TaqMan probe of SEQ ID NO: 65 for 4'CGT.

Torenia DFR

SWB DFR-1158F

5'-AAT GGG ATG CTT CCG ACT TCT-3' (SEQ ID NO: 54)

SWB DFR-1223R

5'-CAG TGG TTT CTG CCA TTG CTT-3' (SEQ ID NO: 55)

SWB DFR-1180T

5'-AGG AAA AAA CAG GCT GAA AA-3' (SEQ ID NO: 56)

Torenia F3H

Torenia F3H-1035F

5'-CAT CGA GCG GTG GTG AAT T-3' (SEQ ID NO: 57)

Torenia F3H-1101R

5'-CTG GCG ATG GGT TTT GAA A-3' (SEQ ID NO: 58)

Torenia F3H-1055T

5'-AAA CAC GAA TAG AAT GTC G-3' (SEQ ID NO: 59)

AS

AmAS-1545F

5'-GAA GAT GAC CTT GCG GTG ATT T-3' (SEQ ID NO: 60)

AmAS-1638R

5'-TTG TCC TCT TCC CCT TTA TAG GTT T-3' (SEQ ID NO: 61)

AmAS-1582T

5'-AGT TCG CCG GGA GTT TCG TGA GTC TG-3' (SEQ ID NO: 62)

4'CGT

AmGTcg12-908F

5'-GGT TGG CCC GCA TTT CA-3' (SEQ ID NO: 63)

AmGTcg12-966R

5'-TAG AAA ACC CTC CGG CAG AA-3' (SEQ ID NO: 64)

AmGTcg12-929T

5'-AGA TGG ACT TAA ATG CG-3' (SEQ ID NO: 65)

As an endogenous control there was used torenia glyceraldehyde phosphate dehydrogenase (GAPDH). The oligoprimers used were SWB GAPDH-794F (5'-GCA TTG AGC AAG ACG TTT GTG-3') (SEQ ID NO: 66) and SWB GAPDH-859R (5'-ACG GGA ACT GTA ACC CCA TTC-3') (SEQ ID NO: 67), and the TaqMan probe used was SWB GAPDH-816T (5'-AGC TTG TGT CGT GGT ACG-3') (SEQ ID NO: 68).

The reaction solution was prepared to a total volume of 50 µl, comprising the cDNA of each transformant line, 1 x TaqMan Universal Master Mix (Applied Biosystems), 100 nM of each oligoprimer and 100 nM of TaqMan probe. Reaction was conducted at 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles of reaction at 95°C for 15 seconds, 60°C for 1 minute, and the production of PCR amplification product was detected in real time. As a result, high expression was confirmed both for introduced AS and 4'CGT in the pSFL201-introduced lines. It was also confirmed that the introduced AS and 4'CGT were both expressed in the pSFL307-introduced lines, and that endogenous DFR mRNA was inhibited to about 10% compared to the stock variety. In the pSFL308-introduced lines, it was confirmed that both the introduced AS and 4'CGT were expressed, while endogenous F3H mRNA was inhibited to about 5% compared to the stock variety.

Example 13. Assay of 4'CGT glucosylating activity on PHC

PHC 4'-glucoside has been confirmed in yellow

Antirrhinum majus flower petals (Sato, T., et al. Plant Sci. 160, 229-236 (2001)), and it is known that AS can produce bracteatin and bracteatin 6-glucoside using PHC and PHC 4'-glucoside as substrate. In order to determine whether or not 4'CGT can catalyze PHC 4'-position glucosylation reaction, the glucosylating activity of 4'CGT for PHC was assayed. Enzyme reaction was conducted according to Example 5-2, using recombinant 4'CGT expressed in *E. coli* and PHC fixed on resin by the same method as Example 3-2 as substrate. The HPLC conditions were as follows.

The column used was a YMC-ODS-A312 (6 mmϕ x 150 mm, YMC Corp.), with a mobile phase of H_2O containing 2% acetic acid as solution A and methanol as solution B, for elution for 22 minutes with a linear concentration gradient from 15% solution B to 40% solution B and holding for 5 minutes with 40% solution B, followed by elution for 14 minutes with a linear concentration gradient from 40% solution B to 62% solution B and holding for 2 minutes with 62% solution B. The flow rate was 1.0 ml/min. Detection was carried out based on absorbance at 360 nm and the absorption spectrum from 220-400 nm using an SPD-M10AP PDA detector (Shimadzu Corp.). Under these conditions, THC elutes at a retention time of 38.2 min, THC 2'-glucoside elutes at 27.7 min, 4'-glucoside elutes at 30.0 min, PHC elutes at 32.4 min and PHC 4'-glucoside elutes at 24.3 min. The PHC glucoside found in the pSPB1768-expressing *E. coli* extract was identified as PHC 4'-glucoside because it eluted at 24.3 min by analysis under these conditions. When PHC was reacted with a crude extract prepared in the same manner from *E. coli* expressing only pQE61 vector, no PHC glucoside was detected. Thus, PHC 4'-glucoside is presumably a product of GT encoded in pSPB1725. These results confirmed that GT encoded in pSPB1725 cDNA exhibits activity of transferring glucose to the 4'-hydroxyl group of PHC. The results described above

demonstrated that 4'CGT catalyzes not only 4'-glucosylation of THC but also 4'-glucosylation of PHC.

Example 14. Functional analysis of 4'CGT and AS using torenia transformants

14-1 Creation of construct

A 2.4 kb 4'CGT expression cassette portion was cut out from the pSFL203 described in Example 7 with PacI, and was inserted at the PacI site of pBINPLUS to obtain plasmid pSFL209. Plasmid pSFL209 expresses 4'CGT alone in plants.

A 2.7 kb F3H expression cassette was cut out from pSFL313 described in Example 9 with AscI, and this was inserted at the AscI site of pBINPLUS to obtain plasmid pSFL210. Plasmid pSFL210 is designed to transcribe double-stranded RNA of the torenia F3H gene in torenia plants, thereby inhibiting expression of F3H.

Separately, for AS, an AS cDNA fragment cut out from pSPB251 described in Example 7 with BamHI and XhoI was inserted at the BamHI and XhoI sites of pSPB120' described in a patent application (P2003-293121), to obtain vector pSPB211 which expresses AS in plants.

14-2 RT-PCR expression analysis of introduced genes and flower color analysis

Plasmids pSFL209, pSFL210 and pSFL211 described in Example 14-1 were introduced into torenia (variety: Summerwave Blue (Suntory Flowers Ltd.)) by the method described in Example 10. The transformation method was based on the method described in Mol. Breeding. 6, 239, (2000). Individuals exhibiting selection marker resistance were selected out. An RNeasy Plant Mini Kit (QIAGEN) was used for extraction of total RNA from buds of each line of the transformants and the stock variety Summerwave Blue, and cDNA was synthesized from 1 μg of the obtained total RNA by reverse transcription using a SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen). RT-PCR reaction was also conducted using ExTaq (TaKaRa) according to the method recommended by the

manufacturer. For amplification of AS mRNA there were used primers AmAS-INSITU-FW (5'-AATTATTTCCCAATGTTCAAAAAT-3') (SEQ ID NO: 44) and AmAS-INSITU-RV (5'-TGGAGCTTTAGGTTTGTGAAA-3') (SEQ ID NO: 45), for amplification of *Antirrhinum majus* 4'CGT mRNA there were used primers KIR-INSITU-FW (5'-ATGGGAGAAGAATACAAGAAAAC-3') (SEQ ID NO: 46) and KIR-INSITU-RV (5'-TCTTACGATAAAACAAACTCA-3') (SEQ ID NO: 47), for amplification of endogenous F3H mRNA there were used primers T.F3H-923F (5'-ATC ATC GAG CGG TGG TGA A-3') (SEQ ID NO: 48) and T.F3H-1339R (5'-TGG CCG ACT AGG CAA TAC AAT-3') (SEQ ID NO: 49), and for amplification of mRNA of GAPDH as the endogenous standard gene there were used primers T.GAPDH-F87 (5'-CCC TTC TGT TTG GTG AAA AGC C-3') (SEQ ID NO: 50) and T.GAPDH-R692 (5'-CCT CGG ATT CCT CCT TGA TAG C-3') (SEQ ID NO: 51). As a result, of the 41 transformant lines obtained among the pSFL209-introduced lines, the introduced *Antirrhinum majus* 4'CGT transcripts were detectable in 37 lines, but no flower color alteration was observed in any of the lines. Of the 44 transformant lines obtained among the pSFL210-introduced lines, significant decrease in the amount of endogenous F3'H transcript was detectable in 37 lines, and these lines exhibited white or mixed violet/white flower colors. In addition, of the 41 transformant lines obtained among the pSPB211-introduced lines, expression of the introduced AS was confirmed in 31 lines, and no flower color alteration was observed in any of the lines.

Pigment analysis was carried out for the pSFL209-introduced lines and pSPB211-introduced lines in which transcripts of the introduced genes were detected, and for the pSFL210-introduced lines that exhibited white flower color. The flower petals of Summerwave Blue and each of the transformants were wetted with 50% acetonitrile containing 0.1% trifluoroacetic acid (TFA) for extraction of the flavonoids, and then AU6-glucoside and anthocyanidins were analyzed by high performance

liquid chromatography (HPLC). Analysis of the anthocyanidins was carried out in the same manner as Example 10. The HPLC conditions were as follows.

First, detection of AU6-glucoside was performed using a SHIM-PACK FC-ODS column (50 x 4.6 mm, Shimadzu Corp.), and for the reverse phase, H_2O containing 0.05% TFA was used as solution A and acetonitrile containing 0.05% TFA was used as solution B. Elution for 3 minutes with a linear concentration gradient from 10%-23% solution B was followed by 17 minutes with 23% solution B, and then elution for 2 minutes with a linear concentration gradient from 23%-80% solution B was followed by 3 minutes with 80% solution B. Elution was then performed for 2 minutes with a concentration gradient from 80%-10% solution B. The flow rate was 0.8 ml/min. Detection was carried out based on absorbance at 360 and 400 nm and the absorption spectrum from 250-500 nm using an SPD-M10AVP PDA detector (Shimadzu Corp.). Under these conditions, THC 4'-glucoside and AU6-glucoside samples elute at retention times of 14.17 min and 6.19 min, respectively. The column used for anthocyanidins was a YMC-ODS-A A312 (6 x 150 mm, YMC Corp.). For the reverse phase there was used a mixture of acetic acid, methanol and H_2O at 60:70:270, maintained for 11 minutes. Detection was carried out based on absorbance at 520 nm and the absorption spectrum from 400-600 nm using an SPD-M10AVP PDA detector (Shimadzu Corp.). Under these conditions, malvidin elutes at a retention time of 9.12 min.

As a result, with the pSFL209-introduced lines, a product matching the retention time and absorption spectrum of THC 4'-glucoside was confirmed to be produced at 0.036-0.762 mg per gram of fresh flower petal weight. Anthocyanidins naturally found in the host were also present.

The anthocyanidin levels in the pSFL210-introduced lines were reduced to about 1% of the anthocyanidin level

in host flower petals.

In the pSPB211-introduced lines, no alteration in flavonoid pigment was observed compared to the host. No product matching the retention times and absorption spectra of aurones was detected in any of the transformants.

This indicated that chalcone glucosides and aurones are not synthesized in plants merely by inhibition of F3H expression and overexpression of AS, but that aurones are synthesized by co-expression of 4'CGT and AS. Overexpression of 4'CGT alone can accumulate chalcone 4'-glucoside, for alteration in flower color.

Example 15. Cloning of 4'CGT cDNA from *Linaria bipartita*

In the same manner as Example 1, RNA extracted from *Linaria bipartita* buds and bloomed flower petals was used to prepare a cDNA library. A library of 8.0×10^5 pfu/ml was obtained from the bud-obtained RNA, while a cDNA library of 1.0×10^6 pfu/ml was obtained using bloomed flower petal cDNA.

Screening was performed using approximately 3.0×10^5 pfu of phage from each library, with 4'CGT cDNA encoded in the *Antirrhinum majus* pSPB1725 described in Example 4 as the probe. The probe labeling, hybridization and subsequent membrane rinsing and detection were carried out in the same manner as Example 2. Nineteen positive clones were finally obtained as a result. When nucleotide sequencing was performed on 8 of the cDNAs having lengths (about 1.5 kb) as expected for cDNA encoding chalcone glucosyltransferase, the 8 clones all had identical sequences and the clone with the longest cDNA was designated as pSFL409. The nucleotide sequence of this cDNA is listed as SEQ ID NO: 69, and the amino acid sequence encoded by it is listed as SEQ ID NO: 70. The amino acid sequence encoded by the pSFL409 cDNA was shown to have high homology with that of *Antirrhinum majus* chalcone 4'-glucosyltransferase. In comparison

with *Antirrhinum majus* chalcone 4'-glucosyltransferase cDNA, however, the amino acid sequence encoded by pSFL409 cDNA was found to be incomplete cDNA lacking about 10 bp from the start methionine. Therefore, a Gene Racer RACE kit (Invitrogen) was used for amplification of the upstream cDNA fragment including the putative start methionine by the 5'RACE method, and this was cloned in pCRII-TOPO vector and designated as plasmid pSFL417. *Linaria bipartita* cDNA including the full length showed 65% sequence identity with *Antirrhinum majus* 4'-glucosyltransferase on the amino acid level.

Example 16. Expression and activity assay of *Linaria bipartita* cDNA in *E. coli*

The full-length *Linaria bipartita* cDNA was introduced at the NcoI and KpnI sites of the *E. coli* expression vector pQE61, and activity of the protein encoded by this *Linaria bipartita* cDNA was analyzed with an *E. coli* expression system. First, for creation of the *E. coli* expression construct, PCR was carried out using pSFL417 as template and primers 417-NcoI (CCCATATATAGCCATGGAAGATACCATCG) (SEQ ID NO: 52) and 409-EcoRI (TAGTGTTGTGGAGTCGGGGGATTTCG) (SEQ ID NO: 53). Thus, an NcoI site was inserted at the start methionine position of pSFL417, and an EcoRI site was inserted at the 3'-end. The products of cleaving this with NcoI/EcoRI and cleaving pSFL409 cDNA with EcoRI/KpnI were cloned at the NcoI/KpnI site of the *E. coli* expression vector pQE61, to obtain an *E. coli* expression construct (pSFL418) comprising the full-length *Linaria bipartita* cDNA.

This was introduced into *E. coli* JM109 and the recombinant protein activity was assayed in the same manner as Example 12. Upon reaction of the pSFL418-possessing *E. coli* extract using THC as substrate, THC 4'-glucoside was detected with a retention time of 30.0 min. In separate reaction of pQE61-possessing *E. coli* extract with THC as a control experiment, absolutely no

THC 4'-glucoside was detected. Glucosylation activity for PHC was also assayed in the same manner as Example 12. As a result, upon reaction of pSFL418-possessing *E. coli* extract, PHC 4'-glucoside was detected with a retention time of 24.3 min. However, no PHC glucoside was detected in the control experiment. These results suggest that *Linaria bipartita* cDNA cloned in SFL418 codes for 4'CGT.